

Cubo OLAP

Para el cubo Olap se hace uso de POWER BI. Inicialmente se cargan los datos desde un .xlsx

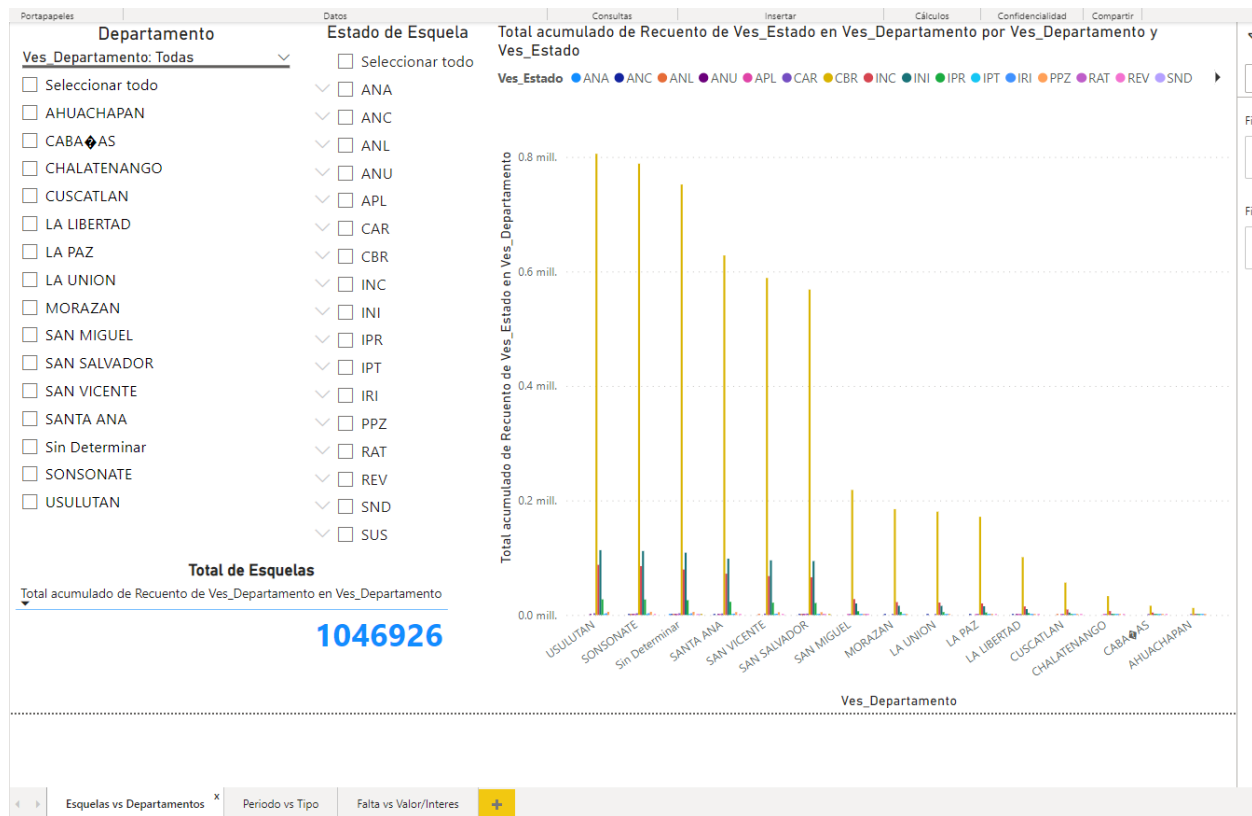
The screenshot shows the Power BI Desktop interface. On the left, the 'Navegador' (Navigator) pane displays the data source 'esquelas_datosmineria.xlsx [3]' with three tables: 'Esquelas' (selected), 'Hoja1', and 'Esquelas1'. The main area shows the 'Esquelas' table with the following columns: 'Ves_Nro_Esquila', 'Ves_Fecha', 'Ves_Tipo_Falta', and 'Ves_Falta_Descripcion'. The table contains 25 rows of data. At the bottom right, there are buttons for 'Cargar', 'Transformar datos', and 'Cancelar'.

| Ves_Nro_Esquila | Ves_Fecha | Ves_Tipo_Falta | Ves_Falta_Descripcion |
|-----------------|-------------------|----------------|--|
| 110637 | 4/6/2003 00:00:00 | 1 | NO UTILIZAR EL CONDUCTOR EL CINTURON DE SEGURIDAD |
| 107742 | 4/6/2003 00:00:00 | 1 | NO UTILIZAR EL CONDUCTOR EL CINTURON DE SEGURIDAD |
| 106645 | 4/6/2003 00:00:00 | 1 | ESTACIONARSE EN CURVAS PUENTES Y TRENCHOS ANGOSTOS |
| 154061 | 4/6/2003 00:00:00 | 1 | BAJAR O SUBIR PASAJEROS EN LUGARES NO PERMITIDOS O ESQUINAS |
| 154102 | 3/6/2003 00:00:00 | 1 | DISPUTARSE LA VIA CON OTRO VEHICULO |
| 119779 | 6/6/2003 00:00:00 | 1 | CIRCULAR VEHICULOS DE TRACCION HUMANA O ANIMAL CON RUEDAS M |
| 106938 | 4/6/2003 00:00:00 | 1 | TRANSPORTAR PERSONAS EN FUNCION COMERCIAL SIN EL PERMISO COF |
| 154457 | 4/6/2003 00:00:00 | 1 | TRANSPORTAR PERSONAS EN FUNCION COMERCIAL SIN EL PERMISO COF |
| 154951 | 6/6/2003 00:00:00 | 2 | NO RESPETAR LAS PARADAS PREVIAMENTE SEALIZADAS POR LA UNIDA |
| 154976 | 6/6/2003 00:00:00 | 2 | NO RESPETAR LAS PARADAS PREVIAMENTE SEALIZADAS POR LA UNIDA |
| 109210 | 6/6/2003 00:00:00 | 2 | TRANSPORTAR PERSONAS EN FUNCION COMERCIAL SIN EL PERMISO COF |
| 109211 | 6/6/2003 00:00:00 | 2 | TRANSPORTAR PERSONAS EN FUNCION COMERCIAL SIN EL PERMISO COF |
| 109207 | 6/6/2003 00:00:00 | 2 | TRANSPORTAR PERSONAS EN FUNCION COMERCIAL SIN EL PERMISO COF |
| 105578 | 9/6/2003 00:00:00 | 1 | CIRCULAR CON VEHICULOS NO MATRICULADOS |
| 109469 | 6/6/2003 00:00:00 | 2 | TRANSPORTAR PERSONAS EN FUNCION COMERCIAL SIN EL PERMISO COF |
| 154826 | 6/6/2003 00:00:00 | 2 | TRANSPORTAR PERSONAS EN FUNCION COMERCIAL SIN EL PERMISO COF |
| 154952 | 6/6/2003 00:00:00 | 2 | CONDUCCION CON LAS PUERTAS ABIERTAS |
| 110968 | 6/6/2003 00:00:00 | 2 | TRANSPORTAR PERSONAS EN FUNCION COMERCIAL SIN EL PERMISO COF |
| 106398 | 6/6/2003 00:00:00 | 2 | CONDUCCION CON LAS PUERTAS ABIERTAS |
| 106075 | 7/6/2003 00:00:00 | 2 | CONDUCCION CON LAS PUERTAS ABIERTAS |
| 107107 | 7/6/2003 00:00:00 | 2 | TRANSPORTAR PERSONAS EN FUNCION COMERCIAL SIN EL PERMISO COF |
| 154977 | 7/6/2003 00:00:00 | 2 | TRANSPORTAR PERSONAS EN FUNCION COMERCIAL SIN EL PERMISO COF |
| 148297 | 7/6/2003 00:00:00 | 1 | CONDUCCION CON LICENCIA CADUCADA |

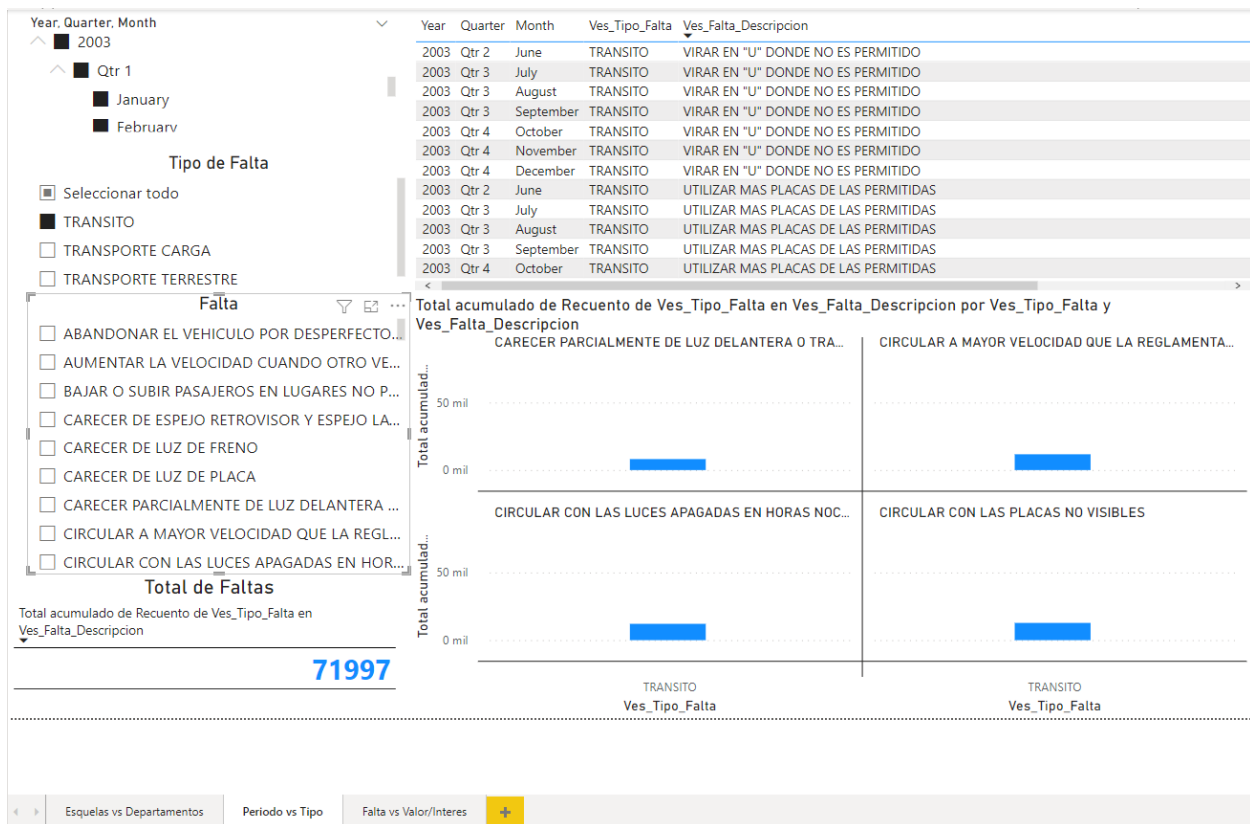
Para analizar los datos desde este cubo se han realizado 3 vistas, cada una representando sus comparaciones: Esquelas Vs. Departamentos, Periodo Vs. Tipo, Falta Vs. Valor/Interés.

Quedando de la siguiente manera:

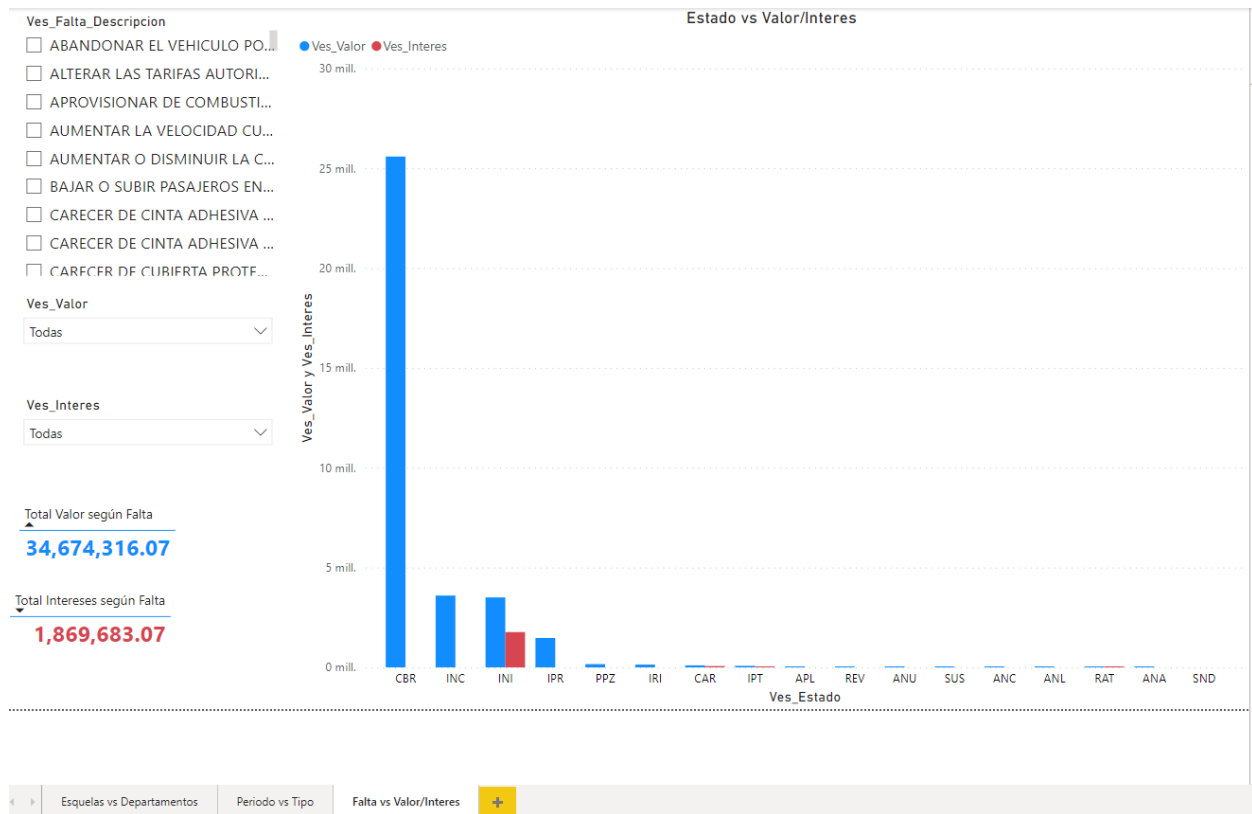
Esquelas Vs. Departamentos, la cual nos permite analizar la cantidad de esquelas por cada uno de los departamentos, distribuidos por el estado de la esquila.



Periodo Vs. Tipo, el cual nos permite analizar los tipos de faltas por periodos (año, cuarto y mes). Esto permite analizar específicamente etapas del año en las que puede haber eventos y comparar si esto produce un alza o baja de esuelas.



Falta Vs. Valor/Interés, permite analizar cada una de las faltas con su valor, así como su acumulado e interés.



Árbol de decisión

Para el árbol de decisión se hace uso de RapidMiner se convertirá el .csv a .xlsx para revisar los datos inicialmente



Vista previa de datos

Esquelas.csv

Origen de archivo: 1252: Europeo occidental (Windows) | Delimitador: Punto y coma | Detección del tipo de datos: Basado en las primeras 200 filas

| Ves_Nro_Esquila | Ves_Fecha | Ves_Tipo_Falta | Ves_Falta_Descripcion | Ves_Departamento | Ves_Est |
|-----------------|-------------------|----------------------|---|------------------|---------|
| 110637 | 4/6/2003 00:00:00 | TRANSITO | NO UTILIZAR EL CONDUCTOR EL CINTURON DE SEGURIDAD | CHALATENANGO | CBR |
| 107742 | 4/6/2003 00:00:00 | TRANSITO | NO UTILIZAR EL CONDUCTOR EL CINTURON DE SEGURIDAD | CHALATENANGO | CBR |
| 106645 | 4/6/2003 00:00:00 | TRANSITO | ESTACIONARSE EN CURVAS PUENTES Y TRECOS ANGO... | SAN SALVADOR | CBR |
| 154061 | 4/6/2003 00:00:00 | TRANSITO | BAJAR O SUBIR PASAJEROS EN LUGARES NO PERMITID... | SAN SALVADOR | CBR |
| 154102 | 3/6/2003 00:00:00 | TRANSITO | DISPUTARSE LA VIA CON OTRO VEHICULO | SAN SALVADOR | |
| 119779 | 6/6/2003 00:00:00 | TRANSITO | CIRCULAR VEHICULOS DE TRACCION HUMANA O ANIM... | SAN SALVADOR | CBR |
| 106938 | 4/6/2003 00:00:00 | TRANSITO | TRANSPORTAR PERSONAS EN FUNCION COMERCIAL SI... | SAN SALVADOR | CBR |
| 154457 | 4/6/2003 00:00:00 | TRANSITO | TRANSPORTAR PERSONAS EN FUNCION COMERCIAL SI... | SAN SALVADOR | CBR |
| 154951 | 6/6/2003 00:00:00 | TRANSPORTE TERRESTRE | NO RESPETAR LAS PARADAS PREVIAMENTE SEÑALIZAD... | SAN SALVADOR | CBR |
| 154976 | 6/6/2003 00:00:00 | TRANSPORTE TERRESTRE | NO RESPETAR LAS PARADAS PREVIAMENTE SEÑALIZAD... | SAN SALVADOR | |
| 109210 | 6/6/2003 00:00:00 | TRANSPORTE TERRESTRE | TRANSPORTAR PERSONAS EN FUNCION COMERCIAL SI... | SAN SALVADOR | CBR |
| 109211 | 6/6/2003 00:00:00 | TRANSPORTE TERRESTRE | TRANSPORTAR PERSONAS EN FUNCION COMERCIAL SI... | SAN SALVADOR | |
| 109207 | 6/6/2003 00:00:00 | TRANSPORTE TERRESTRE | TRANSPORTAR PERSONAS EN FUNCION COMERCIAL SI... | SAN SALVADOR | CBR |
| 105578 | 9/6/2003 00:00:00 | TRANSITO | CIRCULAR CON VEHICULOS NO MATRICULADOS | SAN SALVADOR | CBR |
| 109469 | 6/6/2003 00:00:00 | TRANSPORTE TERRESTRE | TRANSPORTAR PERSONAS EN FUNCION COMERCIAL SI... | SAN SALVADOR | CBR |
| 154826 | 6/6/2003 00:00:00 | TRANSPORTE TERRESTRE | TRANSPORTAR PERSONAS EN FUNCION COMERCIAL SI... | SAN SALVADOR | CBR |
| 154952 | 6/6/2003 00:00:00 | TRANSPORTE TERRESTRE | CONducir con las puertas abiertas | SAN SALVADOR | PPZ |
| 110968 | 6/6/2003 00:00:00 | TRANSPORTE TERRESTRE | TRANSPORTAR PERSONAS EN FUNCION COMERCIAL SI... | SAN SALVADOR | CBR |
| 106398 | 6/6/2003 00:00:00 | TRANSPORTE TERRESTRE | CONducir con las puertas abiertas | SAN SALVADOR | CBR |
| 106075 | 7/6/2003 00:00:00 | TRANSPORTE TERRESTRE | CONducir con las puertas abiertas | SAN SALVADOR | |

Cargar | Transformar datos | Cancelar

Consultas y conexiones

Consultas | Conexiones

1 consulta

Esquelas

Se cargaron 484,531 filas. 16 errores.

Al importar, descubrimos algunos errores:

| Table.ReorderColumns("Consever errores", ("Número de fila", "Ves_Mro_Esuela", "Ves_Fecha", "Ves_Tipo_Falta", "Ves_Falta_Descripcion", "Ves_Departamento", "Ves_Estado", "Ves_Estado_Descripcion", "Ves_Valor", "Ves_Interes", "Ves_Valor", "Ves_Interes")) | | | | | | | | |
|--|---------------------|----------------------|---|------------------|------------|------------------------|-----------|-------------|
| # | Ves_Fecha | Ves_Tipo_Falta | Ves_Falta_Descripcion | Ves_Departamento | Ves_Estado | Ves_Estado_Descripcion | Ves_Valor | Ves_Interes |
| 1 | 19/6/2003 00:00:00 | TRANSITO | NO PORTAR LA LICENCIA DE CONDUCIR | | | | Error | |
| 2 | 27/6/2003 00:00:00 | TRANSITO | NO PORTAR LA LICENCIA DE CONDUCIR | | | | Error | |
| 3 | 24/6/2003 00:00:00 | TRANSITO | NO PORTAR LA LICENCIA DE CONDUCIR | | | | Error | |
| 4 | 3/7/2003 00:00:00 | TRANSITO | NO PORTAR LA LICENCIA DE CONDUCIR | | | | Error | |
| 5 | 30/6/2003 00:00:00 | TRANSITO | NO PORTAR LA LICENCIA DE CONDUCIR | | | | Error | |
| 6 | 3/6/2003 00:00:00 | TRANSITO | NO PORTAR LA LICENCIA DE CONDUCIR | | | | Error | |
| 7 | 6/6/2003 00:00:00 | TRANSITO | NO PORTAR LA LICENCIA DE CONDUCIR | | | | Error | |
| 8 | 26/6/2003 00:00:00 | TRANSITO | NO PORTAR LA LICENCIA DE CONDUCIR | | | | Error | |
| 9 | 10/6/2003 00:00:00 | TRANSITO | NO PORTAR LA LICENCIA DE CONDUCIR | | | | Error | |
| 10 | 25/6/2003 00:00:00 | TRANSITO | NO PORTAR LA LICENCIA DE CONDUCIR | | | | Error | |
| 11 | 22/4/2004 00:00:00 | TRANSPORTE TERRESTRE | TRANSPORTAR PERSONAS EN FUNCION COMERCIAL SIN EL PERMISO C... | SAN MIGUEL | IFT | IMPUESTA | 24.29 | Error |
| 12 | 22/5/2004 00:00:00 | TRANSPORTE TERRESTRE | TRABAJAR EN RUTAS DISTINTAS DE LAS QUE EL VEHICULO TIENE PER... | SAN SALVADOR | IFT | IMPUESTA | 57.14 | Error |
| 13 | 1/7/2004 00:00:00 | TRANSPORTE TERRESTRE | NO PORTAR LOS PERMISOS DE OPERACION DE LINEA | | IFT | IMPUESTA | 57.14 | Error |
| 14 | 16/11/2004 00:00:00 | TRANSPORTE TERRESTRE | NO RESPETAR LAS PARADAS PREVIAMENTE SEÑALIZADAS POR LA UNI... | SAN MIGUEL | IFT | IMPUESTA | 57.14 | Error |
| 15 | 3/12/2004 00:00:00 | TRANSITO | ESTACIONARSE EN ZONA PROHIBIDA O EJE PREFERENCIAL | SAN MIGUEL | IFT | IMPUESTA | 57.14 | Error |
| 16 | 23/11/2004 00:00:00 | TRANSITO | ESTACIONARSE EN ZONA PROHIBIDA O EJE PREFERENCIAL | SAN MIGUEL | IFT | IMPUESTA | 57.14 | Error |
| 17 | 21/10/2006 00:00:00 | TRANSITO | CONDUCCION MOTOCICLETA SIN EL CASCO PROTECTOR | LA UNION | IFT | IMPUESTA | 84.29 | Error |
| 18 | 21/10/2006 00:00:00 | TRANSITO | CONDUCCION SIN ESTAR AUTORIZADO PARA ELLO | LA UNION | IFT | IMPUESTA | 57.14 | Error |
| 19 | 13/7/2007 00:00:00 | TRANSITO | UTILIZAR MAS PLACAS DE LAS PERMITIDAS | ANU | ANULADA | | 11.43 | |

Observamos que algunos campos se encuentran vacíos, vamos a determinar un valor para que no queden null

| Ves_Falta_Descripcion | Ves_Departamento |
|--|------------------|
| NO PORTAR LLANTA DE REPUESTO EN VEHICULOS O CIRCULAR C | SANTA ANA |
| CIRCULAR CON UNA PLACA | SANTA ANA |
| CIRCULAR VEHICULOS DE TRACCION HUMANA O ANIMAL CON R | |
| ESTACIONARSE EN LA ACERA;PARALELO A LA VIA O ATRAVESAD | |
| ESTACIONARSE EN LA ACERA;PARALELO A LA VIA O ATRAVESAD | |
| NO UTILIZAR EL CONDUCTOR EL CINTURON DE SEGURIDAD | |
| NO UTILIZAR EL CONDUCTOR EL CINTURON DE SEGURIDAD | |
| NO UTILIZAR EL CONDUCTOR EL CINTURON DE SEGURIDAD | |
| TRANSPORTAR PERSONAS EN FUNCION COMERCIAL SIN EL PERM | |
| CONDUCCION CON LAS PUERTAS ABIERTAS | |
| CIRCULAR VEHICULOS DE TRACCION HUMANA O ANIMAL CON R | |
| CIRCULAR CON VEHICULOS NO MATRICULADOS | |
| CONDUCCION CON LICENCIA CADUCADA | |
| NEGARSE A PRESENTAR LOS DOCUMENTOS DE TRANSITO A LA A | |
| CIRCULAR VEHICULOS DE TRACCION HUMANA O ANIMAL CON R | |
| CIRCULAR VEHICULOS DE TRACCION HUMANA O ANIMAL CON R | |

Para realizar las modificaciones vamos a apoyarnos de SQL y la BD que ya hemos montado. (Estas mismas modificaciones se realizaron para el Cubo Olap)

Identificamos que el estado cargado, no tiene un valor

| Ves_Estado | Ves_Estado_Descripcion |
|------------|------------------------|
| NULL | CARGADA |
| NULL | CARGADA |
| NULL | CARGADA |
| NULL | CARGADA |
| NULL | CARGADA |
| NULL | CARGADA |
| NULL | CARGADA |
| NULL | CARGADA |
| NULL | CARGADA |
| NULL | CARGADA |
| NULL | CARGADA |
| NULL | CARGADA |
| NULL | CARGADA |
| NULL | CARGADA |
| NULL | CARGADA |
| NULL | CARGADA |
| NULL | CARGADA |
| NULL | CARGADA |
| NULL | CARGADA |
| NULL | CARGADA |
| NULL | CARGADA |

```

update Esquelas
set Ves_Estado = 'CAR'
where Ves_Estado_Descripcion = 'CARGADA'

```

(2859 rows affected)

Completion time: 2022-04-30T10:00:43.7694250-06:00

Se modificaron las columnas

| | Ves_Nro_Esuela | Ves_Fecha | Ves_Tipo_Falta | Ves_Falta_Descripcion | Ves_Departamento | Ves_Estado | Ves_Estado_Descripcion | Ves_Valor | Ves_Interes |
|---|----------------|-----------------------------|----------------------|--|------------------|------------|------------------------|-----------|-------------|
| 1 | 154102 | 2003-06-03 00:00:00.0000000 | TRANSITO | DISPUTARSE LA VIA CON OTRO VEHICULO | SAN SALVADOR | CAR | CARGADA | 57.14 | 36.57 |
| 2 | 154976 | 2003-06-06 00:00:00.0000000 | TRANSPORTE TERRESTRE | NO RESPETAR LAS PARADAS PREVIAMENTE SE#ALIZADAS P... | SAN SALVADOR | CAR | CARGADA | 57.14 | 36.57 |
| 3 | 109211 | 2003-06-06 00:00:00.0000000 | TRANSPORTE TERRESTRE | TRANSPORTAR PERSONAS EN FUNCION COMERCIAL SIN EL ... | SAN SALVADOR | CAR | CARGADA | 34.29 | 21.94 |

Vemos departamentos que no tienen datos

```

select * from Esquelas where Ves_Departamento is null

```

| Ves_Departamento |
|------------------|
| NULL |
| NULL |
| NULL |
| NULL |
| NULL |
| NULL |
| NULL |
| NULL |
| NULL |

```

update Esquelas
set Ves_Departamento = 'Sin Determinar'
where Ves_Departamento is null

```

(146140 rows affected)

Completion time: 2022-04-30T10:13:30.8826350-06:00

Observamos que algunas no tienen valor, o tienen error

```

select distinct Ves_Valor from Esquelas

```

| | Ves_Valor |
|---|--------------|
| 1 | 57.14 |
| 2 | NULL |
| 3 | 34.29 |
| 4 | 171.43 |
| 5 | CARGADA11.43 |
| 6 | 11.43 |
| 7 | 0 |

En este caso vemos que hay un desplazamiento de los datos, que hay que corregir


```
select * from Esquelas where Ves_Valor = 'CARGADA11.43'
```

| | Ves_Nro_Esquila | Ves_Fecha | Ves_Tipo_Falta | Ves_Falta_Descripcion | Ves_Departamento | Ves_Estado | Ves_Estado_Descripcion | Ves_Valor | Ves_Interes |
|----|-----------------|-----------------------------|----------------|-----------------------------------|------------------|------------|------------------------|--------------|-------------|
| 1 | 155133 | 2003-06-19 00:00:00.0000000 | TRANSITO | NO PORTAR LA LICENCIA DE CONDUCIR | Sin Determinar | NULL | NULL | CARGADA11.43 | 7.31 |
| 2 | 154897 | 2003-06-27 00:00:00.0000000 | TRANSITO | NO PORTAR LA LICENCIA DE CONDUCIR | Sin Determinar | NULL | NULL | CARGADA11.43 | 7.31 |
| 3 | 105100 | 2003-06-24 00:00:00.0000000 | TRANSITO | NO PORTAR LA LICENCIA DE CONDUCIR | Sin Determinar | NULL | NULL | CARGADA11.43 | 7.31 |
| 4 | 133876 | 2003-07-03 00:00:00.0000000 | TRANSITO | NO PORTAR LA LICENCIA DE CONDUCIR | Sin Determinar | NULL | NULL | CARGADA11.43 | 7.31 |
| 5 | 121158 | 2003-06-30 00:00:00.0000000 | TRANSITO | NO PORTAR LA LICENCIA DE CONDUCIR | Sin Determinar | NULL | NULL | CARGADA11.43 | 7.31 |
| 6 | 130466 | 2003-08-03 00:00:00.0000000 | TRANSITO | NO PORTAR LA LICENCIA DE CONDUCIR | Sin Determinar | NULL | NULL | CARGADA11.43 | 7.31 |
| 7 | 163321 | 2003-08-06 00:00:00.0000000 | TRANSITO | NO PORTAR LA LICENCIA DE CONDUCIR | Sin Determinar | NULL | NULL | CARGADA11.43 | 7.31 |
| 8 | 158754 | 2003-08-26 00:00:00.0000000 | TRANSITO | NO PORTAR LA LICENCIA DE CONDUCIR | Sin Determinar | NULL | NULL | CARGADA11.43 | 7.31 |
| 9 | 121067 | 2003-08-10 00:00:00.0000000 | TRANSITO | NO PORTAR LA LICENCIA DE CONDUCIR | Sin Determinar | NULL | NULL | CARGADA11.43 | 7.31 |
| 10 | 184830 | 2003-08-25 00:00:00.0000000 | TRANSITO | NO PORTAR LA LICENCIA DE CONDUCIR | Sin Determinar | NULL | NULL | CARGADA11.43 | 7.31 |

```
update Esquelas
set Ves_Estado = 'CAR', Ves_Estado_Descripcion = 'CARGADA', Ves_Valor='11.43'
where Ves_Valor = 'CARGADA11.43'
```

(10 rows affected)

Una vez realizadas todas las modificaciones, se ha generado un BD de Backup, en la cual se sustituyen el nombre de los campos por sus equivalentes en número, quedando de la siguiente manera.

Tipo de falta

| | |
|---|----------------------|
| 1 | TRANSITO |
| 2 | TRANSPORTE TERRESTRE |
| 3 | TRANSPORTE CARGA |

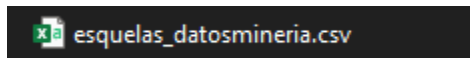
Departamentos

| | |
|----------------|----|
| SAN SALVADOR | 1 |
| CHALATENANGO | 2 |
| MORAZAN | 3 |
| LA PAZ | 4 |
| SONSONATE | 5 |
| SAN MIGUEL | 6 |
| LA LIBERTAD | 7 |
| CUSCATLAN | 8 |
| SAN VICENTE | 9 |
| SANTA ANA | 10 |
| AHUACHAPAN | 11 |
| LA UNION | 12 |
| USULUTAN | 13 |
| CABAÑAS | 14 |
| Sin Determinar | 15 |

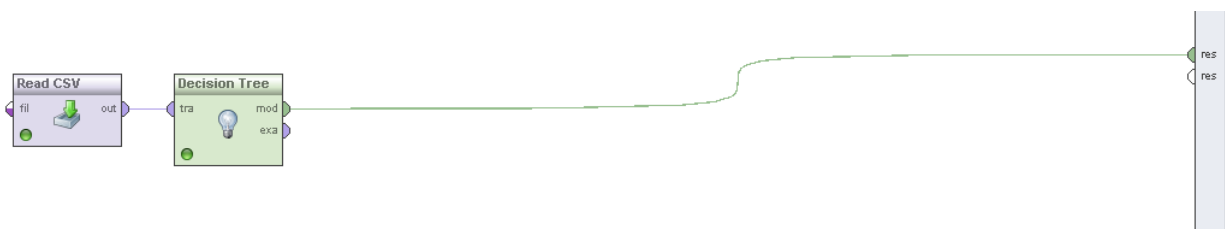
Estado:

| | | |
|----|-----|--------------------------------|
| 1 | SUS | SUSPENSIVO |
| 2 | IPR | IMPROCEDENTE |
| 3 | IPT | IMPUESTA |
| 4 | PPZ | EN PAGOS A PLAZO |
| 5 | CBR | CANCELADA |
| 6 | ANC | ANALISIS COBRADA |
| 7 | INC | INCONSISTENTE |
| 8 | ANU | ANULADA |
| 9 | ANL | ANALISIS |
| 10 | SND | Sin Determinar |
| 11 | INI | PENDIENTE DE PAGO |
| 12 | ANA | ANALISIS APELADA |
| 13 | RAT | RATIFICADA |
| 14 | CAR | CARGADA |
| 15 | REV | REVOCADA |
| 16 | IRI | RECURSO INCONFORMIDAD INICIADO |
| 17 | APL | EN APELACION |

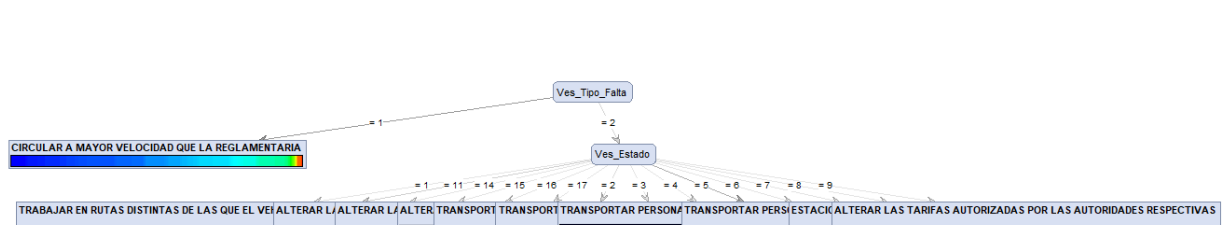
Con estas modificaciones, se exporta un nuevo archivo .csv con el cual trabajaremos en rapid miner



Una vez en rapid miner, se establece el siguiente método:



Dando como resultado el siguiente árbol de decisiones:



Análisis Cubo OLAP y Árbol de Decisiones

Dado el cubo OLAP, árbol y los datos reflejados, se puede realizar el siguiente análisis.

- La esquila mas antigua registrada data de 1968
- El tipo de falta "Transito" contiene la mayor cantidad de faltas: 853,088. Mientras que "Transporte Carga" la menor con: 9134.
- Para el tipo de faltas "Transito", las 4 faltas repetidas son las siguientes:

| Cantidad | Descripción de falta |
|----------|---|
| 53,517 | "No portar licencia de conducir" |
| 45,583 | "Estacionarse en Zona Prohibida o Eje Preferencial" |
| 45,117 | "No utilizar el cinturón de seguridad" |
| 44,559 | "No portar triangulo reflectivo" |

- Para el tipo de faltas "Transporte Carga", las 4 faltas repetidas son las siguientes:

| Cantidad | Descripción de falta |
|----------|--|
| 2,599 | "Circular vehículos que transportan carga con un peso mayor al autorizado" |
| 1,424 | "Carecer de cubierta protectora sobre la carga" |
| 779 | "Carecer de cinta Adhesiva reflejante u ojos de gato reglamentarios" |
| 517 | "Llevar sobre la carga del camión o rastrados personas" |

- Para el tipo de faltas "Transporte terrestre", las 4 faltas repetidas son las siguientes:

| Cantidad | Descripción de falta |
|----------|--|
| 42,600 | "Transportar personas en función comercial sin el permiso correspondiente" |
| 36,130 | "Alterar las tarifas autorizadas por las autoridades" |
| 34,944 | "Estacionarse más tiempo del necesario para subir o bajar" |
| 31,624 | "Conducir con las puertas abiertas" |

- Los tres Departamentos con mayor cantidad de esquelas son:

| Cantidad | Descripción de falta |
|----------|----------------------|
| 485,201 | San Salvador |
| 146,140 | Sin Determinar |
| 82,395 | La Libertad |

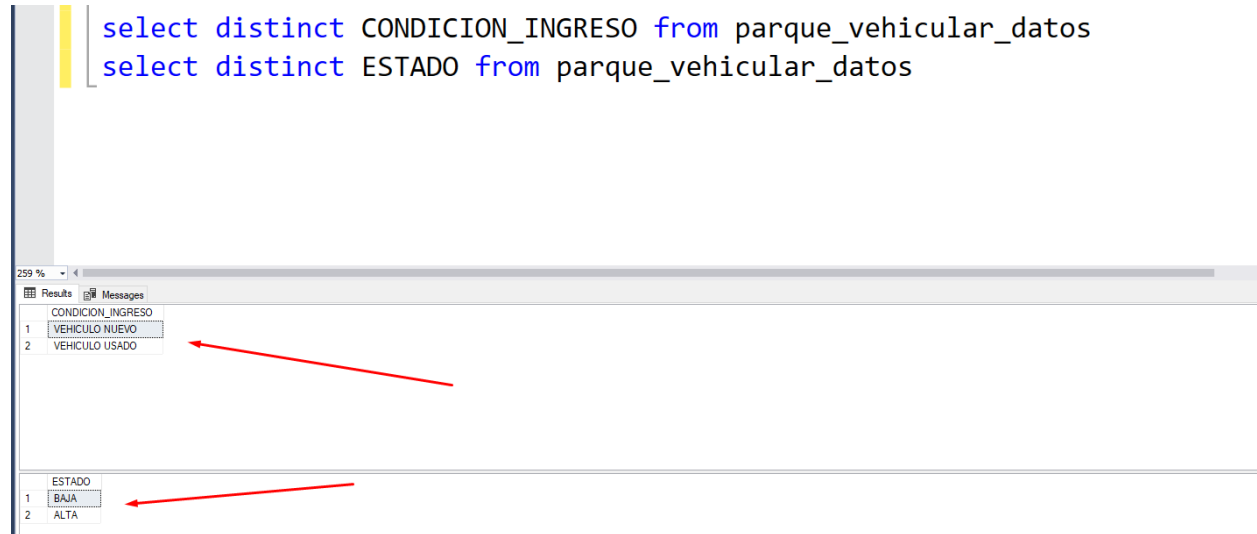
- La esquila que mas se impone tiene un valor de \$57.14, siendo "Muy Grave".

Reglas de asociación

Para este apartado, se utiliza el archivo compartido, el cual se ha cargado en SQL Server para realizar limpieza de datos.

Reglas de asociación establece que los datos deben ser binomiales, por lo cual, como primer punto se identifican los datos que pueden ser tratados como binomiales.

```
select distinct CONDICION_INGRESO from parque_vehicular_datos
select distinct ESTADO from parque_vehicular_datos
```



The screenshot shows the SQL Server Enterprise Manager interface. The top pane displays the results of the first query, 'CONDICION_INGRESO', with two distinct values: 'VEHICULO NUEVO' and 'VEHICULO USADO'. The bottom pane displays the results of the second query, 'ESTADO', with two distinct values: 'BAJA' and 'ALTA'. Red arrows point from the text 'CONDICION_INGRESO' and 'ESTADO' in the code block above to their respective result sets in the screenshot.

| CONDICION_INGRESO |
|-------------------|
| 1 VEHICULO NUEVO |
| 2 VEHICULO USADO |

| ESTADO |
|--------|
| 1 BAJA |
| 2 ALTA |

Para estos datos vamos a establecer el Tipo de Placa como el ID. Se procede a limpiar los datos, estableciendo los datos como 0 y 1.

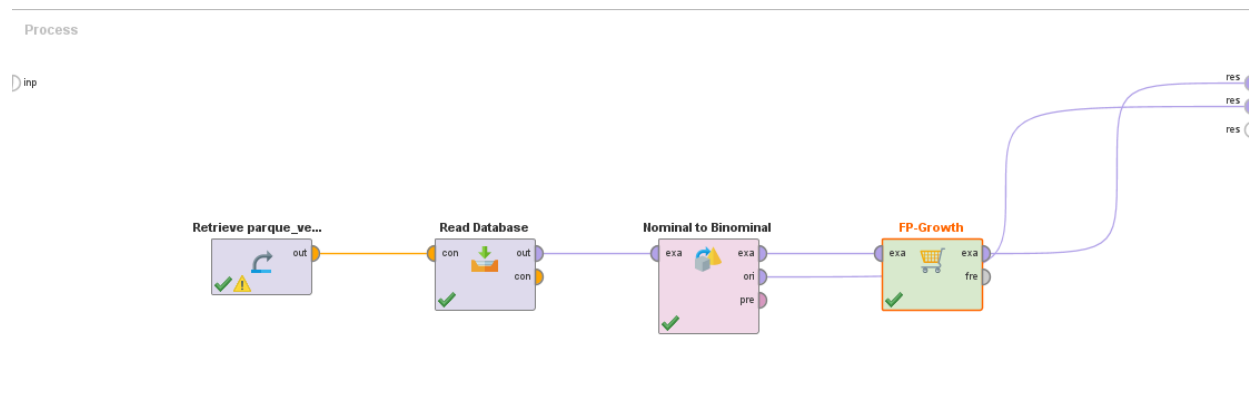
```
select distinct CONDICION_INGRESO from parque_vehicular_datos
update parque_vehicular_datos
set CONDICION_INGRESO = '2'
where CONDICION_INGRESO = 'VEHICULO USADO'

select distinct ESTADO from parque_vehicular_datos
update parque_vehicular_datos
set ESTADO = '2'
where ESTADO = 'ALTA'
```

Se crea una nueva tabla, en la que haremos la carga de los datos que utilizaremos para consulta

| | Column Name | Data Type | Allow Nulls |
|---|---------------------|---------------|-------------------------------------|
| ▶ | TIPO_PLACA_RA | nvarchar(MAX) | <input type="checkbox"/> |
| | CANTIDAD_PUERTAS_RA | float | <input checked="" type="checkbox"/> |
| | COLORES_RA | nvarchar(MAX) | <input checked="" type="checkbox"/> |
| | CLASE_RA | nvarchar(50) | <input checked="" type="checkbox"/> |
| | MARCA_RA | nvarchar(50) | <input checked="" type="checkbox"/> |
| | | | <input type="checkbox"/> |

Una vez trabajados los datos, procedemos a Rapid Miner a crear la estructura de reglas de asociación:



Con un Fp-Growth inicial, tenemos una imagen de los datos de la siguiente manera:

Result History

ExampleSet (Read Database) ExampleSet (Nominal to Binominal)

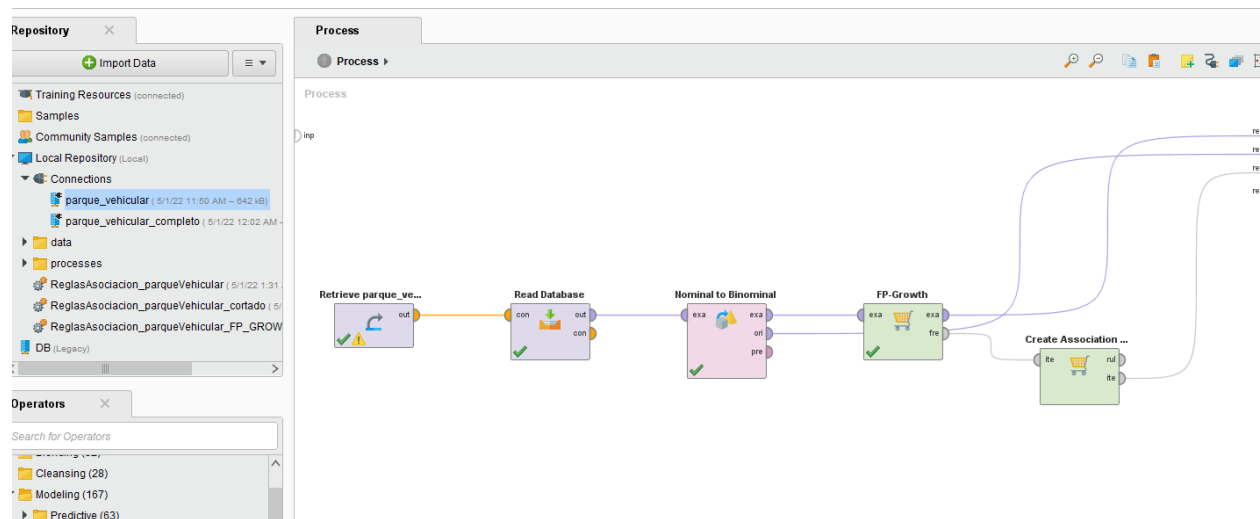
Open in Turbo Prep Auto Model

Filter (1,451,114 / 1,451,114 examples): all

| Row No. | TIPO_PLAC... | CANTIDAD_... | COLORES_RA | CLASE_RA | MARCA_RA |
|---------|--------------|--------------|---------------|-----------|-----------|
| 1 | ALQUILER | 0.00 | AMARILLO | ALQUILER | TOYOTA |
| 2 | PARTICULAR | 0.00 | AMARILLO | AUTOMOVIL | ISUZU |
| 3 | ALQUILER | 0.00 | AMARILLO | ALQUILER | TOYOTA |
| 4 | ALQUILER | 0.00 | AMARILLO | ALQUILER | TOYOTA |
| 5 | PARTICULAR | 0.00 | AMARILLO | AUTOMOVIL | DATSUN |
| 6 | ALQUILER | 0.00 | AMARILLO | ALQUILER | TOYOTA |
| 7 | PARTICULAR | 0.00 | AZUL | AUTOMOVIL | TOYOTA |
| 8 | PARTICULAR | 4.00 | AMARILLO | AUTOMOVIL | CHEVROLET |
| 9 | PARTICULAR | 4.00 | ROJO | AUTOMOVIL | SUBARU |
| 10 | ALQUILER | 0.00 | AMARILLO | ALQUILER | TOYOTA |
| 11 | PARTICULAR | 4.00 | GRIS | AUTOMOVIL | NISSAN |
| 12 | ALQUILER | 0.00 | AMARILLO | ALQUILER | FORD |
| 13 | ALQUILER | 0.00 | AMARILLO | ALQUILER | DODGE |
| 14 | ALQUILER | 0.00 | AMARILLO | ALQUILER | DODGE |
| 15 | PARTICULAR | 0.00 | AMARILLO | AUTOMOVIL | DATSUN |
| 16 | PARTICULAR | 4.00 | NEGRO CON... | AUTOMOVIL | TOYOTA |
| 17 | ALQUILER | 0.00 | AMARILLO | ALQUILER | TOYOTA |
| 18 | ALQUILER | 0.00 | AMARILLO F... | ALQUILER | TOYOTA |
| 19 | PARTICULAR | 4.00 | AMARILLO | AUTOMOVIL | ISUZU |
| 20 | ALQUILER | 0.00 | AMARILLO | ALQUILER | TOYOTA |
| 21 | ALQUILER | 0.00 | AMARILLO | ALQUILER | DATSUN |
| 22 | ALQUILER | 4.00 | AMARILLO F... | ALQUILER | NISSAN |
| 23 | ALQUILER | 0.00 | AMARILLO | ALQUILER | TOYOTA |

ExampleSet (1,451,114 examples, 0 special attributes, 5 regular attributes)

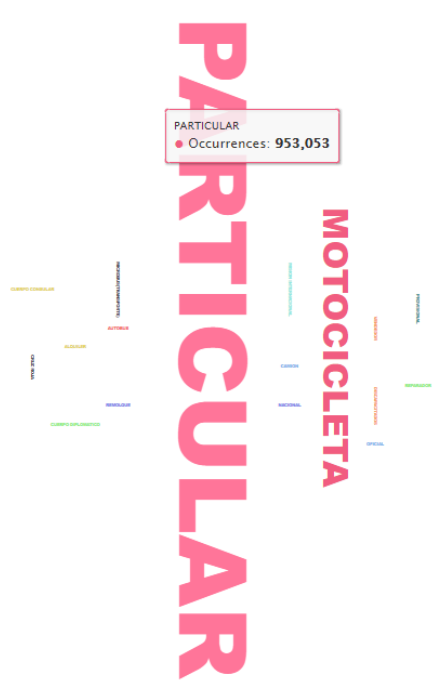
Ahora se procede a crear las reglas de asociación, quedando la siguiente estructura:



Resultando lo siguiente:

| Size | Support | Item 1 | Item 2 | Item 3 |
|------|---------|-----------------------------|----------------------------|----------------------------|
| 1 | 0.657 | TIPO_PLACA_RA = PARTICULAR | | |
| 1 | 0.417 | CLASE_RA = AUTOMOVIL | | |
| 1 | 0.261 | CANTIDAD_PUERTAS_RA = 4.00 | | |
| 1 | 0.242 | TIPO_PLACA_RA = MOTOCICLETA | | |
| 1 | 0.237 | CLASE_RA = MOTOCICLETA | | |
| 2 | 0.413 | TIPO_PLACA_RA = PARTICULAR | CLASE_RA = AUTOMOVIL | |
| 2 | 0.252 | TIPO_PLACA_RA = PARTICULAR | CANTIDAD_PUERTAS_RA = 4.00 | |
| 2 | 0.214 | CLASE_RA = AUTOMOVIL | CANTIDAD_PUERTAS_RA = 4.00 | |
| 2 | 0.237 | TIPO_PLACA_RA = MOTOCICLETA | CLASE_RA = MOTOCICLETA | |
| 3 | 0.212 | TIPO_PLACA_RA = PARTICULAR | CLASE_RA = AUTOMOVIL | CANTIDAD_PUERTAS_RA = 4.00 |

Como resultado podemos ver también los siguientes gráficos



PICK UP

AUTOMOVIL

MOTOCICLETA

MOTOCICLETA

● Occurrences: 344,627

Data from Process

Created by Nominal to Binomial

Rows: 1,451,114

Columns: 5

Last Change: None

| TIPO_PLACA_RA Category | CANTIDAD_PUERTAS_RA Category | COLORES_RA Category | CLASE_RA Category | MARCA_RA Category |
|---------------------------|---------------------------------|------------------------|----------------------|----------------------|
| ALQUILER | 0.00 | AMARILLO | ALQUILER | TOYOTA |
| PARTICULAR | 0.00 | AMARILLO | AUTOMOVIL | ISUZU |
| ALQUILER | 0.00 | AMARILLO | ALQUILER | TOYOTA |
| ALQUILER | 0.00 | AMARILLO | ALQUILER | TOYOTA |
| PARTICULAR | 0.00 | AMARILLO | AUTOMOVIL | DATSUM |
| ALQUILER | 0.00 | AMARILLO | ALQUILER | TOYOTA |
| PARTICULAR | 0.00 | AZUL | AUTOMOVIL | TOYOTA |
| PARTICULAR | 4.00 | AMARILLO | AUTOMOVIL | CHEVROLET |
| PARTICULAR | 4.00 | ROJO | AUTOMOVIL | SUBARU |
| ALQUILER | 0.00 | AMARILLO | ALQUILER | TOYOTA |
| PARTICULAR | 4.00 | GRIS | AUTOMOVIL | NISSAN |
| ALQUILER | 0.00 | AMARILLO | ALQUILER | FORD |
| ALQUILER | 0.00 | AMARILLO | ALQUILER | DODGE |
| ALQUILER | 0.00 | AMARILLO | ALQUILER | DODGE |
| PARTICULAR | 0.00 | AMARILLO | AUTOMOVIL | DATSUM |
| PARTICULAR | 4.00 | NEGRO CON FRASIAS | AUTOMOVIL | TOYOTA |
| ALQUILER | 0.00 | AMARILLO | ALQUILER | TOYOTA |

1,451,114 rows - 5 columns (2 months)

Creating HTML5 chart

Auto Model

Load Data

Select Task

Prepare Target

Select Inputs

Model Types

Results

RESTART

BACK

NEXT

Predict

Clusters

Outliers

Want to predict the values of a column?

Want to identify groups in your data?

Want to detect outliers in your data?

| TIPO_PLACA_RA | CANTIDAD_PUERTAS_RA | COLORES_RA | CLASE_RA | MARCA_RA |
|---------------|---------------------|-------------------|-----------|-----------|
| Category | Category | Category | Category | Category |
| PARTICULAR | 0.00 | AMARILLO | AUTOMOVIL | ISUZU |
| ALQUILER | 0.00 | AMARILLO | ALQUILER | TOYOTA |
| ALQUILER | 0.00 | AMARILLO | ALQUILER | TOYOTA |
| PARTICULAR | 0.00 | AMARILLO | AUTOMOVIL | DATSUN |
| ALQUILER | 0.00 | AMARILLO | ALQUILER | TOYOTA |
| PARTICULAR | 0.00 | AZUL | AUTOMOVIL | TOYOTA |
| PARTICULAR | 4.00 | AMARILLO | AUTOMOVIL | CHEVROLET |
| PARTICULAR | 4.00 | ROJO | AUTOMOVIL | SUBARU |
| ALQUILER | 0.00 | AMARILLO | ALQUILER | TOYOTA |
| PARTICULAR | 4.00 | GRIS | AUTOMOVIL | NISSAN |
| ALQUILER | 0.00 | AMARILLO | ALQUILER | FORD |
| ALQUILER | 0.00 | AMARILLO | ALQUILER | DODGE |
| ALQUILER | 0.00 | AMARILLO | ALQUILER | DODGE |
| PARTICULAR | 0.00 | AMARILLO | AUTOMOVIL | DATSUN |
| PARTICULAR | 4.00 | NEGRO CON FRANJAS | AUTOMOVIL | TOYOTA |

Load Data

Select Task

Prepare Target

Select Inputs

Model Types

Results

RESTART

BACK

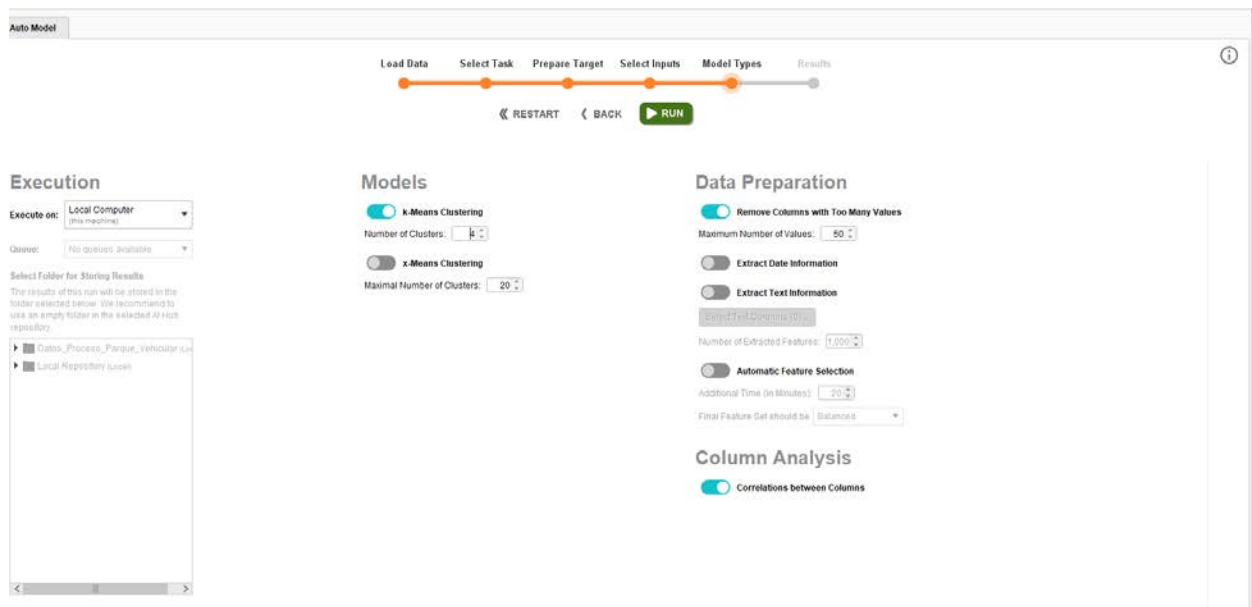
NEXT

Selected: 5 / Total: 5

Select All

Deselect All

| Selected | Status | Quality | Name | Correlation | ID:ness | Stability | Missing | Test:ness |
|-------------------------------------|--------|---|---------------------|-------------|---------|-----------|---------|-----------|
| <input checked="" type="checkbox"/> | ● | <div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div> | TIPO_PLACA_RA | ? | 0.00% | 66.53% | 0.00% | 4.50% |
| <input checked="" type="checkbox"/> | ● | <div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div> | CANTIDAD_PUERTAS_RA | ? | 0.00% | 43.10% | 39.32% | 1.78% |
| <input checked="" type="checkbox"/> | ● | <div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div> | COLORES_RA | ? | 0.42% | 53.34% | 0.01% | 15.66% |
| <input checked="" type="checkbox"/> | ● | <div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div> | CLASE_RA | ? | 0.00% | 41.62% | 0.00% | 13.56% |
| <input checked="" type="checkbox"/> | ● | <div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div> | MARCA_RA | ? | 0.04% | 17.93% | 0.00% | 4.53% |

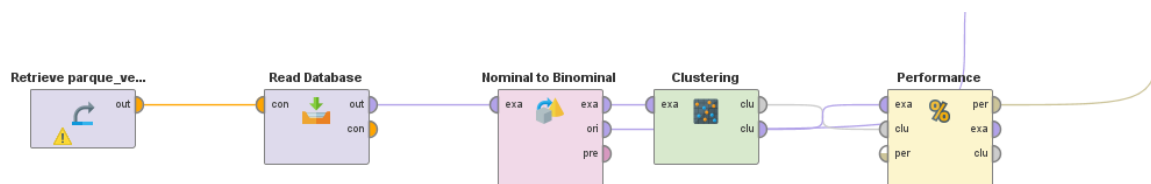


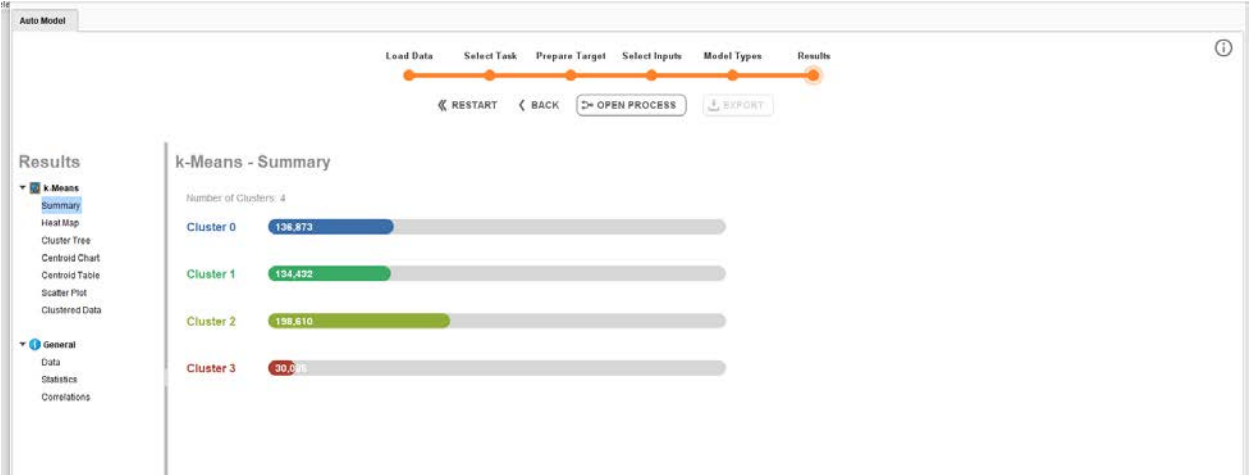
Agrupamiento con K-Means

Para los K-MEANS haremos uso de una tercer tabla creada

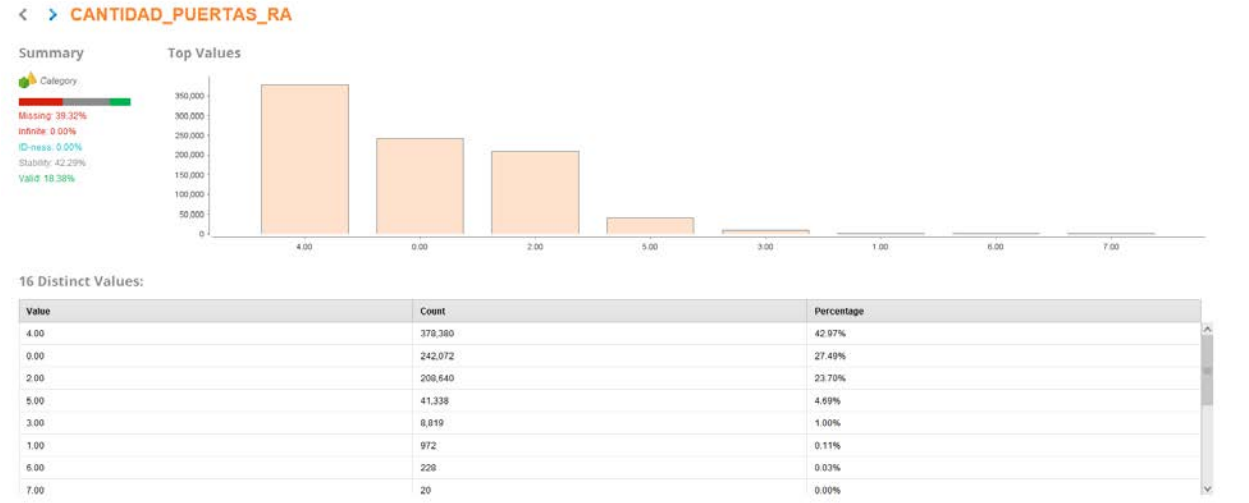
| Column Name | Data Type | Allow Nulls |
|-------------------------|--------------|-------------------------------------|
| ANIO_FABRICACION_KN | nvarchar(50) | <input checked="" type="checkbox"/> |
| ANIO_INGRESO_KN | float | <input checked="" type="checkbox"/> |
| MES_INGRESO_KN | nvarchar(50) | <input checked="" type="checkbox"/> |
| PERTENENCIA_KN | nvarchar(50) | <input checked="" type="checkbox"/> |
| COMBUSTIBLE_KN | nvarchar(50) | <input checked="" type="checkbox"/> |
| ADUANA_KN | nvarchar(50) | <input checked="" type="checkbox"/> |
| PROPIETARIO_DEPARTA... | nvarchar(50) | <input checked="" type="checkbox"/> |
| PROPIETARIO_MUNICIPI... | nvarchar(50) | <input checked="" type="checkbox"/> |
| ESTADO | nvarchar(50) | <input checked="" type="checkbox"/> |

Se establece la siguiente estructura en rapid miner





Statistics



Fase 2

Para esta fase, se realiza un archivo global en PowerBI, el cual contiene todos los datos necesarios para realizar el análisis.

Reporte PowerBI: <https://app.powerbi.com/groups/me/reports/436e1639-aec1-4e7c-adfb-2d6aebc6e220?ctid=f9afe020-14e8-4555-b638-b98f896aa94b>